DAU PRESS
TECHNICAL REPORT, TR 1-00

DEFENSE ACQUISITION UNIVERSITY
DEFENSE SYSTEMS MANAGEMENT COLLEGE
FORT BELVOIR, VIRGINIA

ACQUISITION TREND METRICS IN THE DEPARTMENT OF DEFENSE

MAJOR WILLIAM J. SWANK, USAF
DR. PAUL A. ALFIERI
CHARLES K. GAILEY III
RAYMOND W. REIG

October 2000

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **OCT 2000**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Acquisition Trend Metrics in the Department of Defense**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defense Acquisition University Defense Systems Management College Fort Belvoir, VA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**29**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ACQUISITION TREND METRICS IN THE DEPARTMENT OF DEFENSE

Major William J. Swank, USAF
Dr. Paul A. Alfieri
Charles K. Gailey III
Raymond W. Reig

October 2000

DEFENSE SYSTEMS MANAGEMENT COLLEGE
PUBLISHED BY THE DEFENSE ACQUISITION UNIVERSITY PRESS
Fort Belvoir, Virginia 22060-5565

EXECUTIVE SUMMARY

For five years the Office of the Secretary of Defense (OSD) has funded the Engineering and Manufacturing Development (EMD) Performance Trends Research Project at the Defense Systems Management College (DSMC). This research analyzes the cost, schedule and performance success of Acquisition Category I programs during the EMD phase of development. This Executive Summary contains the major conclusions and recommendations of the research effort to date. Supporting rationale and details are contained in subsequent sections of this report.

Major Conclusions

1. The average cost overrun of programs within EMD for the period 1980 to 1996 was 40 percent; the average schedule overrun was 62 percent. The median cost overrun for the same period was 20 percent; the median schedule overrun was 46 percent. The average performance success rating was 3.72 out of 5. The average duration of EMD for programs between 1980 and 1996 was 7.7 years. Figures are available by year groups and by Service.

2. Schedule slips, particularly of early test events, become the best early indicators of trouble.

3. For programs that must submit an exception SAR before they are through 50 percent of their actual EMD duration, it is most probable they will have a schedule overrun of more than two years.

4. Using only Research, Development, Test and Engineering (RDT&E) costs during EMD provides a comparable and possibly better analysis of the true EMD cost growth percentage than using the combined RDT&E and Procurement costs when interested in managerial effectiveness.

5. The complexity of these acquisition programs makes it difficult to attribute a particular outcome to any set of two (or even three) program characteristics.

6. The OSD/Office of Acquisition Resources and Analysis maintains a single trend metric, "Major Defense Acquisition Program Cost Growth." This is a required macro-economic indicator but does not reflect trends in program execution.

7. The DSMC Research project is the only system of metrics that uniquely combines performance data (operational test results) with general management data (cost and schedule) within EMD.

Recommendations

1. The quarterly Defense Acquisition Executive Summary process and report is the Department of Defense (DoD) method to address potential problems (early warning). It appears to be working well and it should continue to be used by the top acquisition managers within DoD.

2. There appears to be a lack of sufficient **trend** data available in DoD to determine if acquisition management is improving. This research project, or equivalent, could be used as an additional trend metric and should be continued.

3. Next year is the first year sufficient programs that have benefited from the Acquisition Reform (AR) initiatives will have completed EMD and their EMD success can be compared to pre-AR programs using the data in the DSMC spreadsheet.

4. The public perceives DoD acquisition as inefficient. However, there is data showing that DoD acquisition is relatively good when compared to other public and private major developments. Since perception is as important as reality, it is recommended that the relative efficiency of DoD acquisition be advertised.

5. The majority of programs have schedule overruns of less than 85 percent according to schedule/overrun data, while fewer than a third of the programs had schedule overruns of more than 125 percent. This appears to represent a significant difference between the two groups. Further analyses should consider treating these groups separately.

6. Multi-variant analysis of program outcomes should be performed, to isolate significant factors or combinations of factors in success or failure.

TABLE OF CONTENTS

LIST OF FIGURES

ACRONYM LIST

ACAT	Acquisition Category
A/C	Aircraft
ADCAP	Advanced Capability
AFOTEC	Air Force Operational Test and Evaluation Center
AMRAAM	Advanced Medium Range Air-to-Air Missile
APAM	Anti-Personnel Anti-Materiel
APB	Acquisition Program Baseline
AR	Acquisition Reform
AR&A	Office of Acquisition Resources and Analysis (OSD)
ASPJ	Airborne Self-Protection Jammer
ATACMS	Army Tactical Missile System
C/SCSC	Cost/Schedule Control Systems Criteria
DAB	Defense Acquisition Board
DAES	Defense Acquisition Executive Summary
DoD	Department of Defense
DOT&E	Director, Operational Test and Evaluation
DSB	Defense Science Board
DSMC	Defense Systems Management College
EMD	Engineering and Manufacturing Development
FMTV	Family of Medium Tactical Vehicles
IOT&E	Initial Operational Test and Evaluation
IPT	Integrated Product Team
JROC	Joint Requirements Oversight Council
JSOW	Joint Standoff Weapon
JSTARS	Joint Surveillance Target Attack Radar System
JTIDS	Joint Tactical Information Distribution System
LRIP	Low Rate Initial Production
LFT&E	Live Fire Test and Evaluation
MDAP	Major Defense Acquisition Program
M/S	Milestone
MM III GRP	Minuteman III Guidance Replacement Program
PM	Program Manager
PMO	Program Management Office
OA	Operational Assessment
ORD	Operational Requirements Document
OSD	Office of the Secretary of Defense
OT	Operational Testing
RDT&E	Research, Development, Test and Evaluation
SAR	Selected Acquisition Report
SMART-T	Secure, Mobile Antijam, Reliable, Tactical Terminal
T&E	Test and Evaluation
TEMP	Test and Evaluation Master Plan
USD(A&T)	Office of the Deputy Under Secretary of Defense for Acquisition and Technology

1. Background

In 1994 the Science Advisor to the DOT&E, or Director, Operational Test and Evaluation, asked the Defense Systems Management College Test and Evaluation Department (DSMC T&E) to research the relationship, if any, between the number of test articles used in the Engineering and Manufacturing Development (EMD) phase of a program acquisition, and the success of that program in EMD. Since determining the number of test articles required is a difficult trade-off among several factors, the DOT&E wanted to know if data was available that would help in making this important decision. Indications were that factual metrics relating to the subject were nonexistent.

Unless specifically detailed in this report, all metrics and author's conclusions are detailed in DSMC Press Technical Report TR 1-99. (See Bibliography.) Figure 1, also contained in TR 1-99, shows the average cost and schedule overruns as ratios of planned to actual figures. To obtain overrun percentages, subtract 1 from the values on the axes. The average actual duration of EMD for any program, for any year in which that program's EMD phase ended, can be found in the database section of TR 1-99.

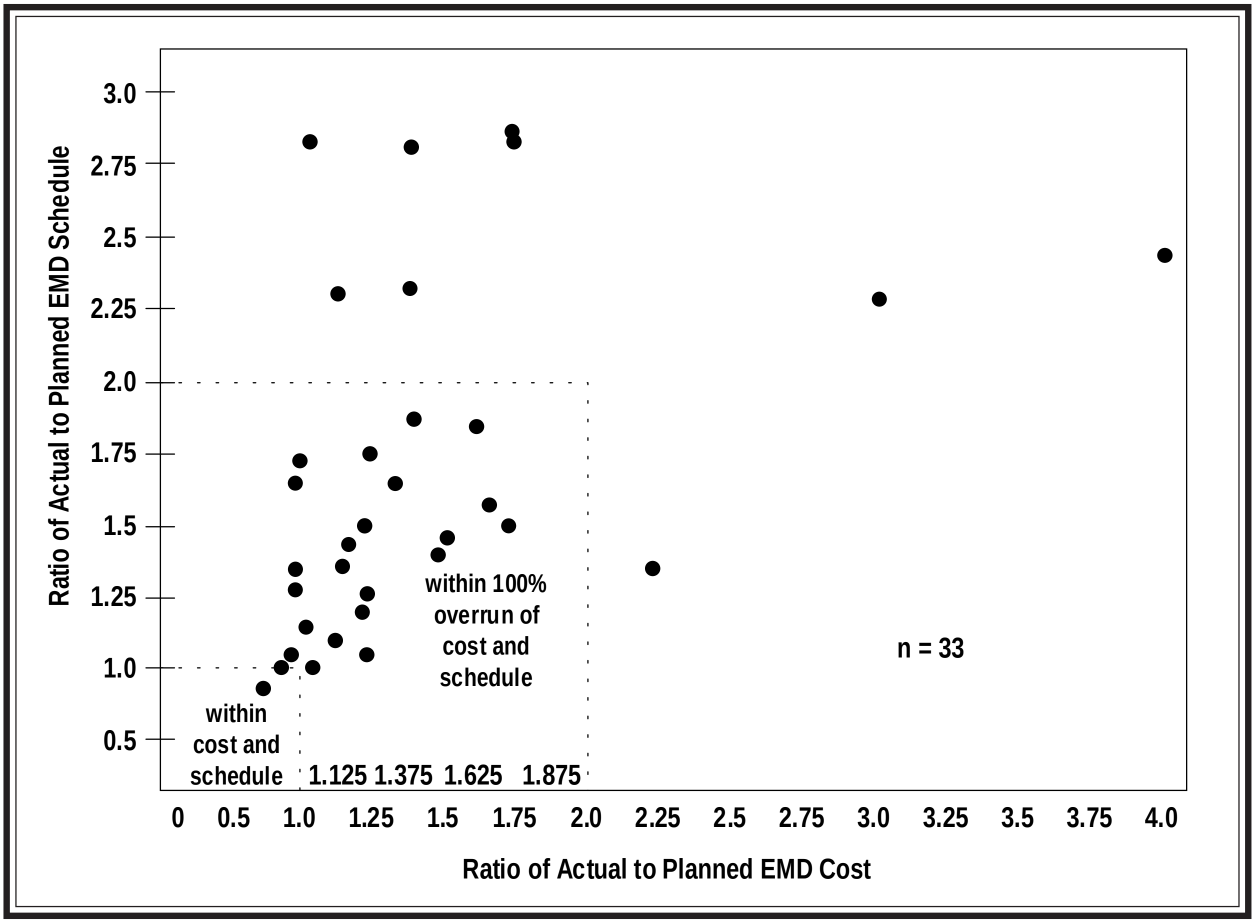


Figure 1. Ratio of Actual to Planned EMD Cost and Schedule

2. DoD Relative Cost Growth

Figure 2 indicates the cost growth (overruns) of Department of Defense (DoD) major weapon systems compared to other complex public and private sector projects. It shows that DoD cost overruns are relatively small when compared to the majority of the other projects shown. Although Figure 2 represents 1985, the close similarity of the 38 percent cost growth of 1985 when compared to the 40 percent cost growth of this research suggests the data is still valid. An updated Figure 2 every ten years or so might prove interesting and valuable. The database of the current research could be used to provide the DoD with updated cost data. A similar chart could be created for schedule slips.

3. Leading Indicators

Recently the research team was asked to identify leading indicators of a program's success or problems within EMD, by a review of the research data available. This would consist of data obtained from the Selected Acquisition Reports (SAR), the Operational Testing (OT) reports, and the DOT&E independent evaluation of the Service OT reports, together with the sum of the data contained in the project spreadsheet.

Neither the Service OT report nor the DOT&E evaluation can provide leading indicators because they occur near the end of the EMD phase. Likewise the data within the database is historical and for any program the data is not entered until the program completes EMD.

The SAR report is a thoroughly reviewed, legally required annual program report (quarterly by exception) that is sent to Congress. Its primary purpose is to be the historically correct record from program inception (Milestone (M/S) I) through 90 percent delivered or expended. The SARs have been analyzed several different ways, and two potential early indicators were identified. (See pages 4 and 9.)

4. Cost, Schedule and Performance as Early Warnings

Cost and schedule problems (overruns) are usually the effect of some other program perturbation, most often shortfalls in performance. Performance, then, would be the best place to look for indications that a program is in trouble.

Unfortunately, there are difficulties obtaining performance data with acceptable confidence early in the program, due to limited numbers of test trials (see page 10). Also, within the program documentation (Acquisition Program Baseline (APB), SAR, and the Test and Evaluation Master Plan (TEMP)), performance is listed by threshold value, objective value, current estimated value, and demonstrated value to date. In the early stages of EMD it is perfectly normal for the demonstrated performance to be below the final threshold requirement. But with EMD durations averaging 8 years or more, it is not reasonable to expect a program manager to say early on that the eventual threshold requirement will not be met.

Cost status is not a good place to look for early indicators of trouble. As mentioned, cost growth is usually a consequence of some other occurrence and not a lead event. Also, cost status lends itself

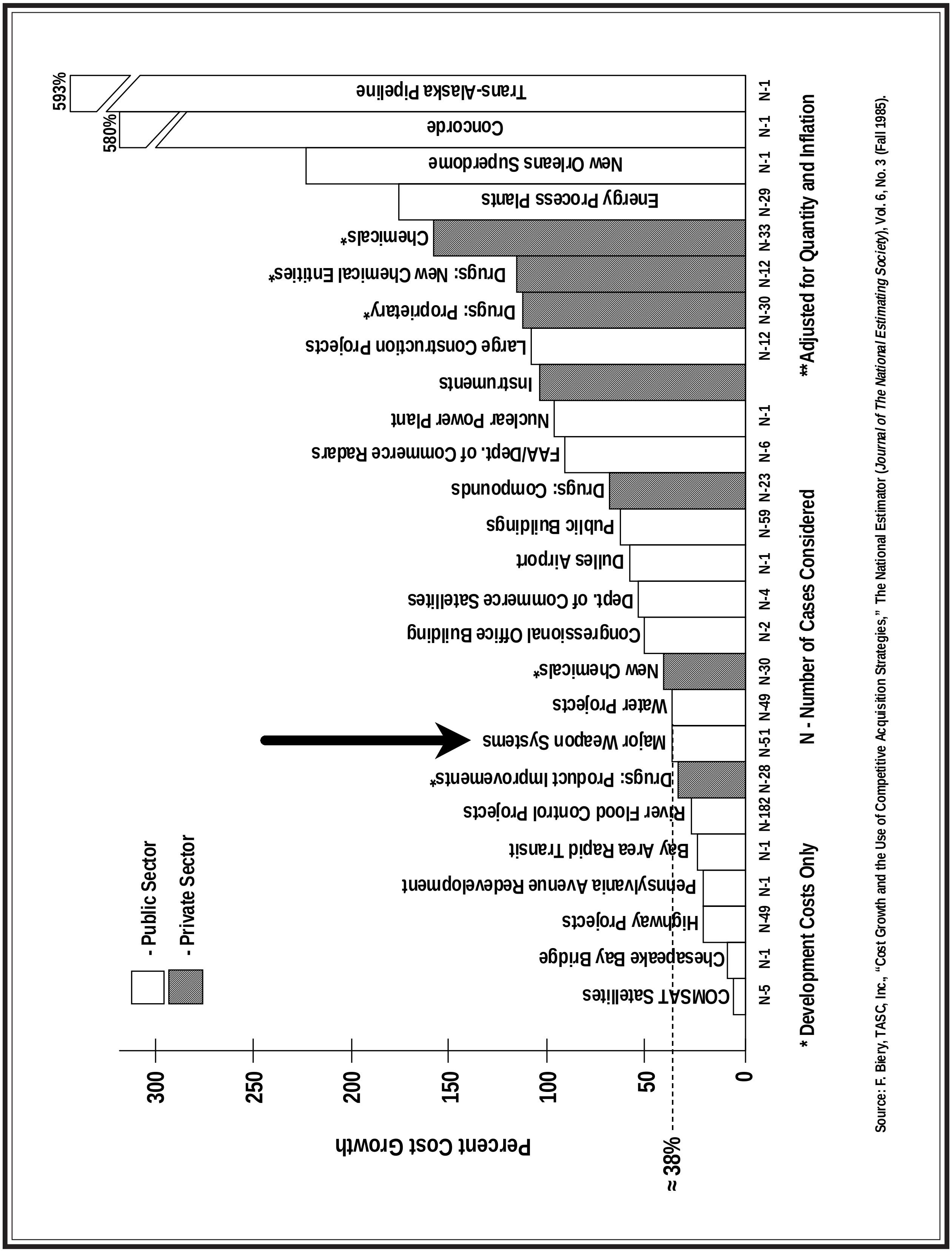


Figure 2. Cost Growth in Major Projects (TASC)

more to manipulation than does schedule or performance. For example, the program can be descoped by reducing logistics support or test programs. Activities can also be postponed into the production phase. The research did, however, uncover some interesting aspects of program costs and cost reporting. (See page 8.)

Schedule status, then, becomes the best early indicator of trouble. To provide early indication, the milestones section of the APB, SAR, DAES, and TEMP must contain sufficient early test start and stop milestones. **Given sufficient detailed milestones in program documentation, a slip in early test milestones is the best indicator of a program heading into trouble.**

5. Sole Trend Metric

The research team believes the Office of the Secretary of Defense (OSD) office most involved with the analysis of the status of DoD acquisition management is the OSD/AR&A (Office of Acquisition Resources and Analysis). That office provides OSD involvement and oversight into the APB, DAES and SAR procedures, as well as other responsibilities. Figures 3 and 4 describe the sole trend metric maintained by the office. This is a macro-economic metric designed to record and control overall funding levels for Major Defense Acquisition Program (MDAP), but does not reflect trends in program execution.

To better gauge improvements in DoD acquisition management over time, we recommend that the DAES process get extensive use, that detailed test milestones be required within SARs and other program documentation, and that additional trend metrics be used (such as the DSMC EMD Performance Trends Research, or equivalent).

- In July 1984 the Congressional Budget Office (CBO) first published statistic "Annual Rate of Cost Change" using Selected Acquisition Report (SAR) data reflecting President's Budgets (FY 80-85)
- DoD added data points (1986-2000)

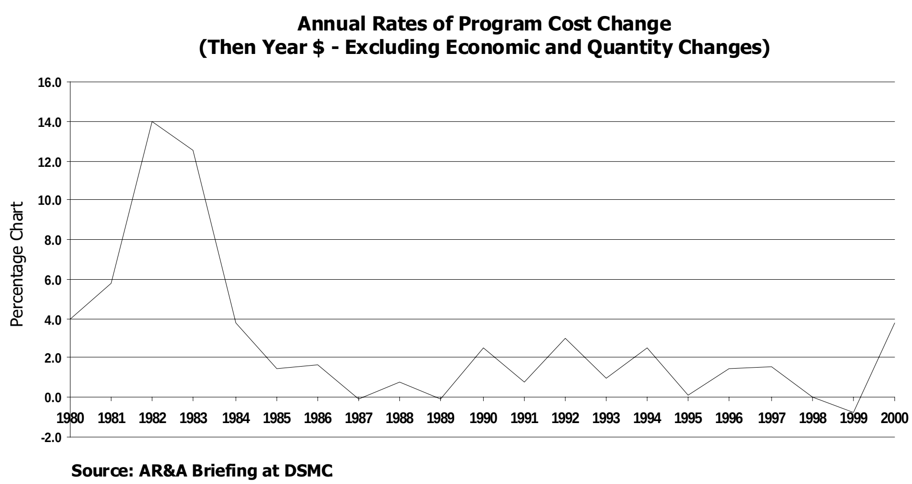
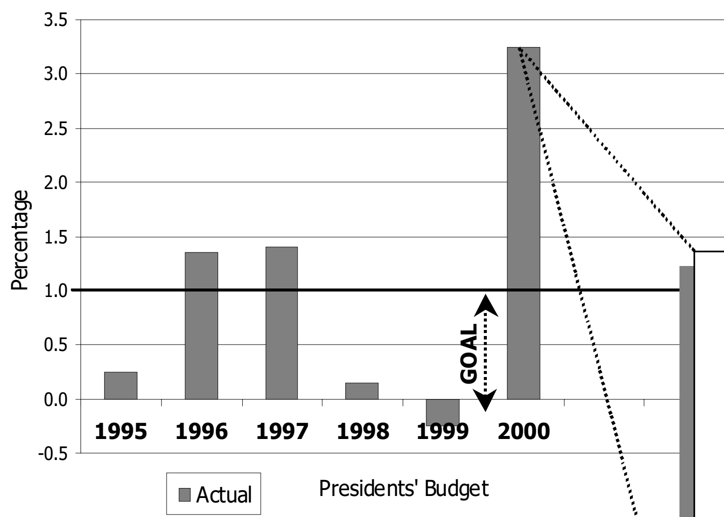


Figure 3. MDAP Cost Growth Metric - Background

Annual Cost Change



- <u>Outcome Metric</u>:
  % change of MDAP total acquisition cost
- <u>Goal</u>:
  Maintain no greater than 1% annually
- <u>Goal Measurement</u>: *Annually, with PPBS projections for the POM and BES*

FY 00 President's Budget Actual

	<u>% Change</u>	<u>Primary Drivers +/-%</u>
Army	+6.3	ATIRCM (+23.7), BAT (+21.8), FMTV (+17.4), LB APACHE (+9.7)
Navy	+1.5	F/A-18 E/F (+4.1), JSF (+6.5), JSOW (-24.3), SSN 774 (+4.4)
Air Force	+1.5	C-17A (+2.2), NAVSTAR (+8.7), SBIRS (+19.4), TITAN IV (-5.1)
BMDO	+23.8	NAVY AREA (+13.0+, NMD (+43.4), PAC-3 (+17.9), THAAD (+22.4)
Overall	+3.1%	

Source: AR&A Briefing at DSMC.

Figure 4. MDAP Cost Growth Metric - Results for Past 6 Years

6. Mean vs. Median

The mean cost (n=33) and schedule (n=34) overruns for programs whose EMD phase ended between 1980 and 1996 were 40 percent in cost and 62 percent in schedule. The median overruns were 20 percent in cost and 46 percent in schedule.

Cost overrun. The large difference between mean and median mostly reflects the presence in the data set of two extreme values: 309 percent for Joint Tactical Information Distribution System (JTIDS) and 201 percent for Joint Surveillance Target Attack Radar System (JSTARS). These two alone account for a 9 percent increase in the mean due to the presence of JTIDS [(309-31)/32] and a 5 percent increase in the mean due to the presence of JSTARS [(210-35)/32]. One commonly used method to describe a data set without the skewing due to one or two extreme numbers is to use the median. In the case of cost overruns, the median of 20 percent better describes the general set of programs than does the mean. This is clearly demonstrated by the mean of 31 programs (26 percent) and mean of 30 programs (23 percent), which shows the means of the more homogeneous populations approaching the median of the general population.

Schedule overrun. The difference between mean and median is smaller, 62 percent vs. 46 percent. This is because there are no extreme values in the data set as there are in the cost overruns. The two largest schedule overruns were 175 percent (Maverick) and 172 percent (Family of Medium Tactical Vehicles). Together, these two overruns added only 7 percent to the mean. In fact, there were four programs with cost overruns between 167 percent and 175 percent: the above two plus Airborne Self-Protected Jammer (ASPJ) and Tomahawk. In this case, it can be argued that four such figures, representing more than 10 percent of the programs, are too many to be non-representative of the data set. In such a case, the mean is a valid statistic to describe the set.

There is at least one other descriptive possibility for this type of data set: the possibility of a bi-modal or even tri-modal population. This involves examining the data to see if there are two or three separate groups within the set. If there are, analysis would have to explore possible reasons for such groupings among the development programs. Schedule overruns appear to form at least three populations: -20 percent to 40 percent (n=15), 40 percent to 80 percent (n=10), and 125 percent to 175 percent (n=9). The first two may form the same population.

The nature of bi-modal populations is that any mean taken for the whole group is irrelevant; each sub-group has to be analyzed separately. In fact, for schedule overrun, it may be productive to look at just two groups: <80 percent (n=26)[mean=32, median=31.5] and >125 percent (n=9)[mean=149, median=136]. Of course, it may be productive to examine in what ways the groups differ, other than in the primary variable of schedule overrun. In the case of schedule overrun, the >125 percent group consists of JTIDS, Maverick, Tomahawk, Advanced Medium Range Air to Air Missile (AMRAAM), B-1B, ASPJ, Mk 48 Advanced Capability (ADCAP), and JSTARS. The significant analytical question is whether these programs had anything in common, which was not present in the other programs.

7. The DAES Process and Report

The Defense Acquisition Executive Summary (DAES) process and report are recommended as the principal method for effective oversight of DoD acquisition management and are briefly described on Figures 5 and 6.

- Monthly *management process*;
- Based on *quarterly reports* from Program Managers;
- Addresses:
  - ♦ Program execution status against the Acquisition Program Baseline;
    - – Cost, Schedule, and Performance Goals.
  - ♦ Potential problems (early warning).

Source: AR&A Briefing at DSMC.

Figure 5. Defense Acquisition Executive Summary (DAES)

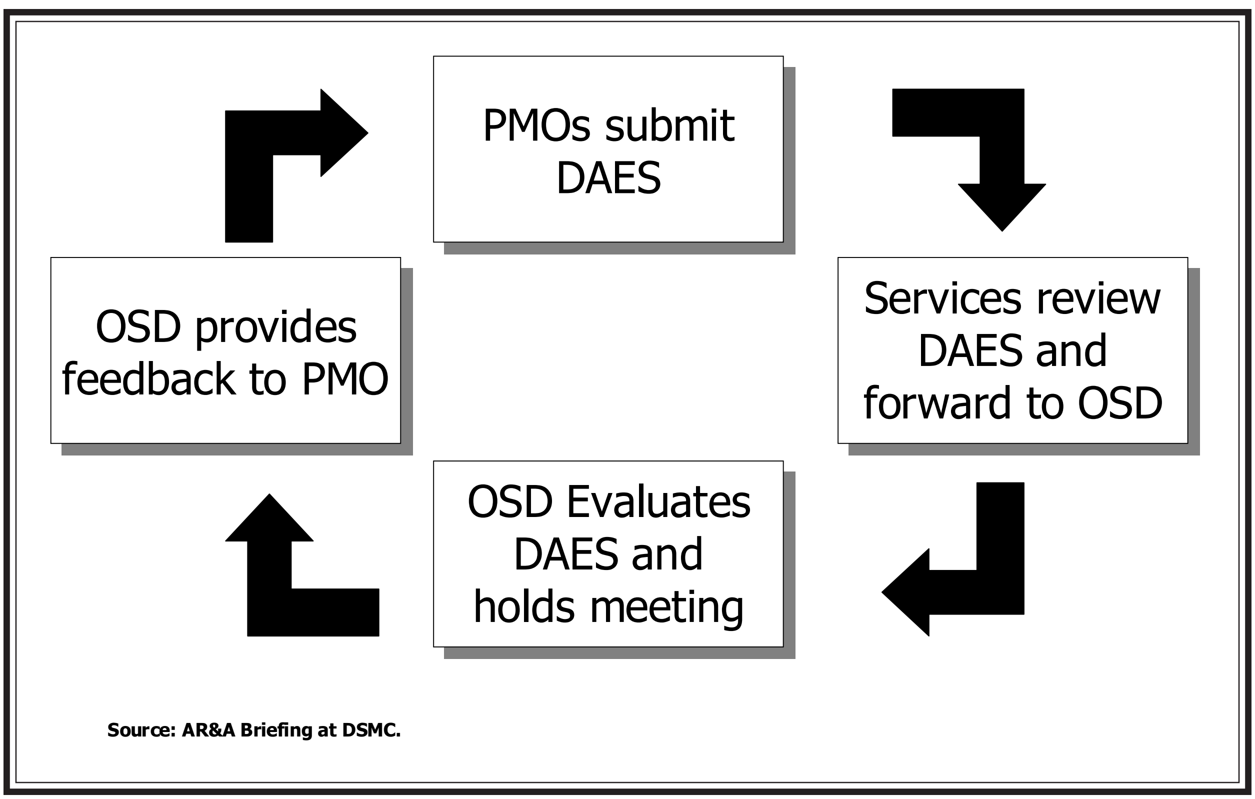


Figure 6. DAES Process

8. Aspects of EMD Costs

Consistent with the idea that all data should be that which are available to OSD-level decision makers, cost data for all systems were collected from the system SARs. As the period of interest was EMD, the Research, Development, Test and Evaluation (RDT&E) and Procurement costs were collected for the period from M/S II through M/S III. Planned costs were those reported in the SAR following M/S II; actual costs were those reported in the SAR following M/S III.

In theory, one could go to the program SAR and retrieve, rather directly, the planned and actual costs of EMD. In practice, the retrieval of costs from the SAR is not simple. For one thing, costs are reported in several places in the SAR: for example, sections 11, 13, 14, and 16. For another, costs are reported in both Base Year and Then Year dollars; the inflators for each appropriation are not always explicit. And yet another, the costs are reported by fiscal year. EMD need not start or end in conjunction with fiscal years. This leaves the analyst with the problem of determining how much of a particular fiscal year's costs should be allocated to EMD for the spreadsheet.

<u>Base Year vs. Then Year.</u> In principle, the use of Base Year figures will remove the effect of inflation on the figures, making it easier to perform analyses. In the case of the EMD spreadsheet, the costs are used as a ratio of planned to actual. For most programs, the same inflators will be used in both sets of figures, thereby yielding the same results from either Base or Then Year figures. For those programs in which actual EMD fiscal years were significantly longer than planned, when inflation might have had an effect on the ratios, use of either figure will still identify a significant overrun. In addition, several programs changed their Base Year for SAR reporting during EMD. This creates additional difficulties in using Base Year figures for the spreadsheet. For all the above reasons, this analysis used Then Year costs.

<u>SAR Sections.</u> Sections 11, 13 and 14 do not separate costs by fiscal year. It is not possible to determine costs for EMD alone from these figures. Section 16 is the sole portion of the SAR that records costs by fiscal year. Start of EMD (M/S II) found in section 9 determines the earliest fiscal year to use. If M/S II is in the middle of a fiscal year, costs can be pro-rated month for month. M/S III determines the end of EMD. If M/S III is in the middle of a fiscal year, the end of Initial Operational Test and Evaluation (IOT&E) can be used to determine the end of EMD expenditures. With the above adjustments, section 16 figures were used for the SAR cost data.

<u>RDT&E Costs.</u> For some programs it is difficult to determine what RDT&E costs during the period of EMD actually were incurred in development of the program. In particular, when the program includes several models, or includes preplanned product improvements, the program can incur RDT&E costs supporting these other projects during the period that the basic development is in EMD. An excellent example of this is the (JTIDS) program, in which class 2H/2M development was initiated before class 2 completed M/S III. SAR data can not isolate the costs of the basic program, except in those cases where the program office identified the parts of the project separately in the SAR. Additionally, for programs with preplanned product improvements, total cost figures (such as are in sections 11, 13 or 14) will not define the basic development. In those instances where cost definition was not clear, best estimates of allocation to the basic development, year by year, were made.

Procurement Costs. The SARs do not separately identify the cost of Low Rate Initial Production, (LRIP). The best approximation is to use all procurement costs for those years before M/S III. This is predicated on the legal constraint against using full rate production funds before that decision. Procurement costs also have the same separation problem as discussed for RDT&E above. For programs with several models, the same solution was used.

RDT&E vs. Procurement. There remains a significant problem in determining the costs to be used in calculating EMD cost overrun. For any program with more than a few LRIP items, the procurement costs in EMD will be significantly (in some cases overwhelmingly) larger than RDT&E costs. This results in the program being scored for EMD cost overrun predominantly by procurement costs. The problem is that a large number of LRIP items may indicate a very bad program (they need to continue EMD to keep the line warm while the program keeps trying to pass IOT&E, for example). A large number of LRIP items may also indicate a very good program (testing in EMD has been so successful the Services demand the item in the field before M/S III for example). Thus, overwhelming RDT&E with Procurement costs produce equivocal data. For this reason, EMD costs for RDT&E and Procurement were entered separately in the spreadsheet, with RDT&E costs alone being used for the EMD cost overrun calculation. Analysis of procurement costs may still be performed, if desired, since the data is present in the spreadsheet. It should be noted that a sensitivity analysis was performed on the data from the systems reported in 1995. The result was that the use of RDT&E costs alone produced the same EMD cost success factor, within a score of 1, in 13 out of 14 programs.

9. Timing of First Exception SAR

The SAR process requires that an exception SAR be submitted to OSD within 90 days after the program office becomes aware of either a cost or schedule breach of the APB. An exception SAR must also be provided if there is a Nunn-McCurdy unit cost breach. The only performance breach known to the research team is if a program fails to meet its exit criteria. Meeting exit criteria requirements is necessary for a program to exit one stage of the development program and proceed into the next phase.

The hypothesis: if the first exception SAR had to be submitted early in the program, it might be used as an early indicator of a program having difficulty. Fifteen program SARs were reviewed to determine when the first exception SAR was submitted on each program. The timing was expressed as a percentage of total time the program was in EMD. It ranged from 9 percent through to 99 percent. The first exception SAR for many programs was submitted more than 50 percent through EMD and thus could not be used as an early indicator.

Figure 7 shows that for programs whose first exception SAR was before the program was 50 percent through EMD, the eventual schedule success at the end of EMD was a low success rating of 1 for each of the eight programs involved. **Apparently, submission of a program's first exception SAR before EMD is 50 percent complete is an indicator of eventual schedule difficulties.**

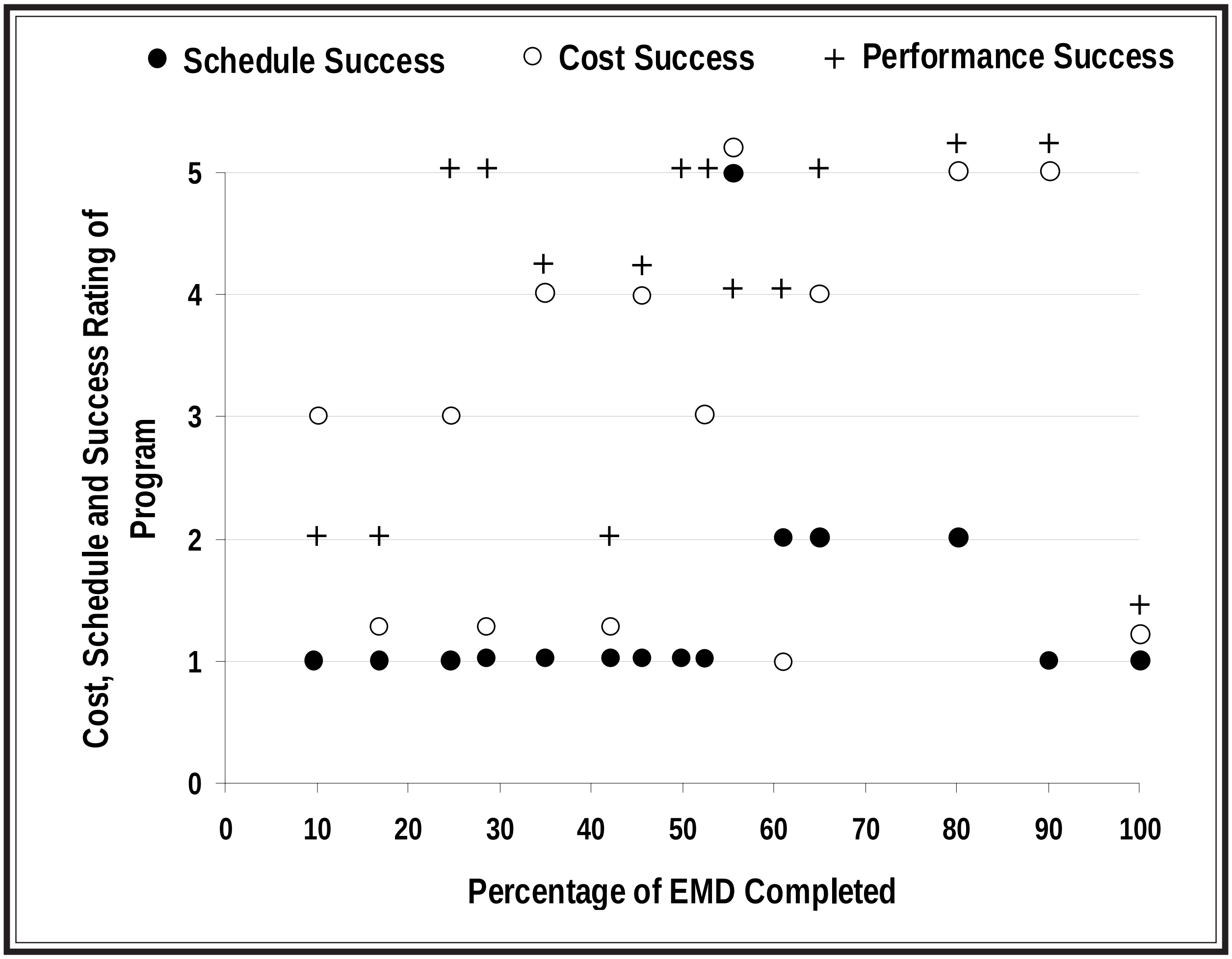


Figure 7. Time (Percentage) into EMD of First Exception SAR

Schedule Success Rating Table

Success Rating	Schedule Overrun
5	On Schedule
4	Less than 12 months
3	Less than 18 months
2	Less than 24 months
1	More than two years

10. Test Type Errors

Hypothesis testing is used to determine whether or not a statement about a population performance should or should not be rejected. That is, an inference concerning the population performance is made based on a sample of that population. The null hypothesis (H_o) expresses an assumption that the component, subsystem or system under test meets the requirement. The alternative hypothesis (H_a) is that the component, subsystem or system doesn't meet the requirement. Four

possible outcomes are possible. 1) H_o is not rejected when the actual population meets the requirement, 2) H_o is rejected when the actual population doesn't meet the requirement, 3) H_o is rejected when the actual population meets the requirement, and 4) H_o is not rejected when the actual population doesn't meet the requirement. The first two outcomes are correct conclusions and the latter two represent errors. The first error (H_o is rejected when the actual population meets the requirement) is a Type I or producer's error. The second (H_o is not rejected when the actual population doesn't meet the requirement) is a Type II or consumer's error.

Within the acquisition and operational communities, when a Type I error occurs the developers must take action to correct deficiencies to meet requirements that are already met. When a Type II error occurs, the component, subsystem or system that doesn't meet requirements is advanced to the next stage of development or fielded. The deficiencies are then uncovered when they are more expensive to fix or when the user needs the capability. As the risk of committing a Type I error increases, the risk of committing a Type II error decreases and visa-versa.

Based on Figure 7, programs with Exception SARs had a noticeably better performance success rating then cost or schedule ratings. This indicates a proclivity to ensure system performance meets requirements before fielding the system. Almost all of these systems entered and finished IOT&E later then planned, but the overall inferred performance indicates operationally effective and suitable systems.

These trends suggest a historical development process and operational test and evaluation community willing to accept increased producer's error to reduce the risk of fielding systems that don't meet operational requirements.

11. Review of the Technical Section of the SAR

The technical section of the SARs during EMD for the JSTARS and AMRAAM were reviewed to determine if any leading indicators of program success or trouble could be found. This involved reading nine JSTAR and six AMRAAM SARs rendered during EMD.

For JSTARS, the technical section started with 6 technical and 6 operational parameters. These were increased in number and later decreased in number. Many explanatory notes were necessarily included to define the conditions of evaluating parameters. Demonstrated values appeared late in EMD, with many parameters to be determined (TBD).

For AMRAAM, there were 7 operational characteristics and technical characteristics in the areas of weight, length, Reliability and Maintainability, and Built-in-Test. There was a robust listing of technical criteria and good notes. There were very few test milestones listed.

The conclusion: it is not possible to obtain leading indicators of a program problem in the technical section of the SAR. Most Acquisition Category (ACAT) I programs are so complex that not even a very good engineer analyst, not in the program office, could follow the technical evolution of the system.

Interestingly, in the JSTAR program, where there were good, detailed test milestones, the first indication of program difficulties was a slip of Air Force and Army test start dates mentioned in the second program EMD SAR.

12. Acquisition Reform Impact

Figure 8 shows that there have been 4 SAR programs which have completed EMD and which had most of their EMD phase after implementation of Acquisition Reform measures. These programs were Army Tactical Missile Systems (ATACMS) Block 1A (Anti-Personnel Anti-Materiel (APAM)), Joint Standoff Weapon (JSOW), Secure Mobile Antijam Reliable Tactical Terminal (SMART-T) and Minuteman (MM) III Guidance Replacement Program (GRP).

Program	**Cost**		**Schedule**		**Performance (OSD Overall)**
	Success	**% overrun**	**Success**	**% overrun**	**Success**
ATACMS	4	5	3	16	4.3
JSOW	4	10	4	3	3.3
SMART-T	4	27	4	4	1.3
MMIII GRP	3	43	1	31	4
Mean	3.75	21	3	34	3.2
Median *	4	19	3.5	10	3.65
For comparison					
All programs 1980-1996					
Mean	3.3	40	2.5	62	3.8
Median *		20		46	

What is striking about the post-AR group is the excellent Schedule success. However, in looking deeper into this result, another factor may be significant. Two of the four post-AR programs were Mods. Modification programs can be expected to have better cost and schedule results than new development programs, and the data demonstrate this.

	Cost % overrun	Schedule % overrun
All programs 1980-1996		
Mods		
Mean	3	45
Median *	9	4
Non-Mods		
Mean	49	68
Median *	35	52

Therefore, the excellent results for post-AR programs might simply be because half of them were modification programs. Examination of programs that have completed EMD under AR policies, to determine the effectiveness of those policies, must wait for more data, particularly a larger study population of such programs.

*For a comparison of the use of medians rather than means to describe these populations, see page 6.

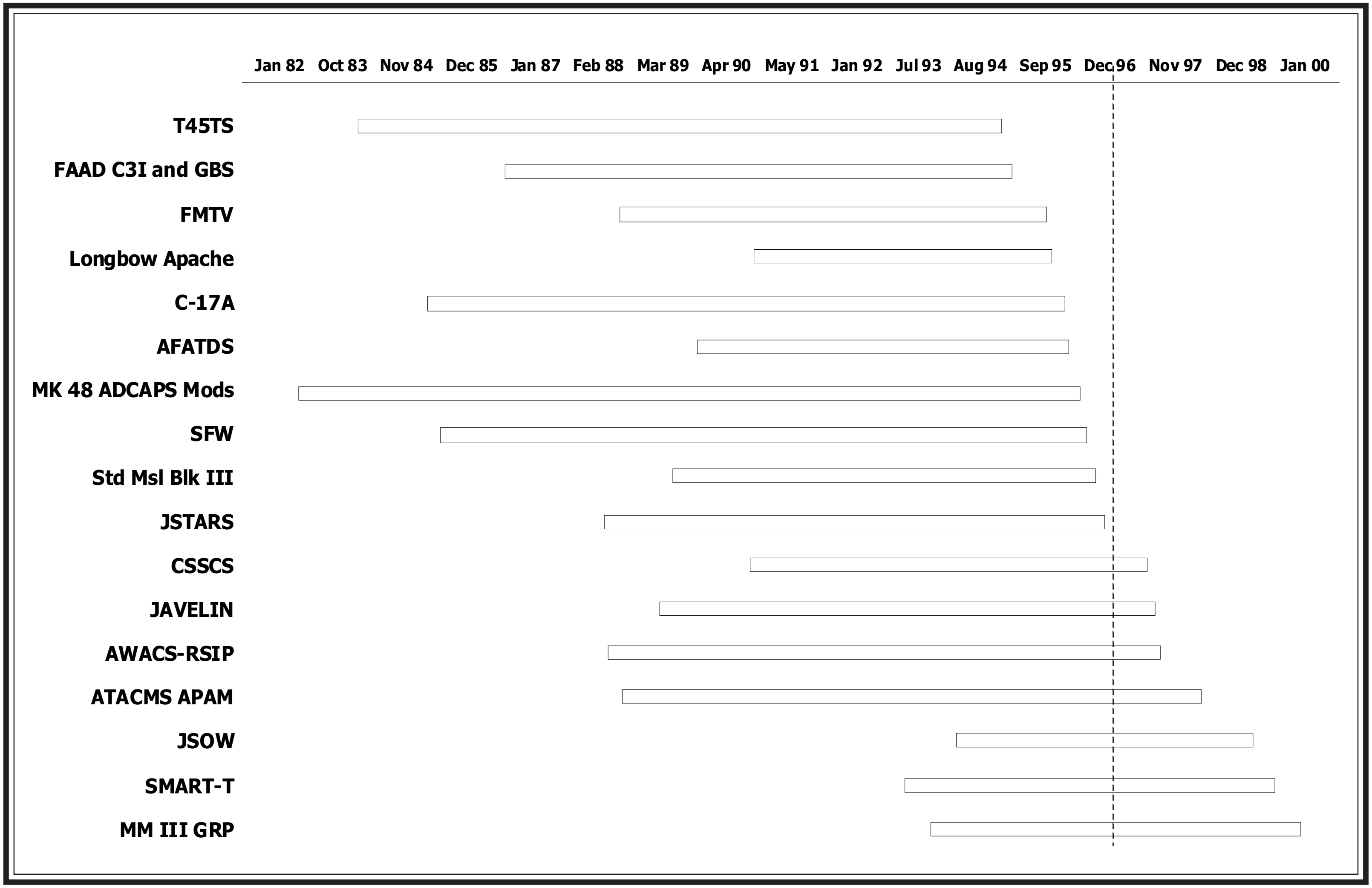


Figure 8. Effect of Acquisition Reform on Program EMD

13. Use of OA during EMD

Programs within the database were reviewed to determine if a program used an OA and if that had any effect on the cost, schedule, or performance success of that program. Performance success is the DOT&E overall success rating for the program.

OA USED

		Success Rating	
Program No.	Cost	Schedule	Performance
11	5	5	4
37	5	2	5
30	4	5	3
18	5	1	5
20	5	2	2
13	4	1	1
36	3	1	5
17	2	1	3
41	1	1	2
n = 9 Average	3.8	2.1	3.3

OA NOT USED

		Success Rating	
Program No.	Cost	Schedule	Performance
1	3	4	5
5	4	4	4
9	5	4	3
40	5	1	5
7	4	1	5
19	4	1	5
31	4	1	4
10	1	1	4
22	3	1	2
3	1	2	1
23	1	1	1
n = 11 Average	3.2	1.9	3.5

Initial Conclusions

The use of OAs had a positive result for cost and schedule success and a negative result for performance. Schedule and performance success differences were close to each other. Cost success using OAs seems to have made the greatest difference. In all cases the data is broadly scattered.

Since these results are counter-intuitive, more analysis and more data are required.

14. Program Characterization

During the research effort it was necessary to read every SAR written on a program during EMD. This being so, it was thought that a short one- or two-page summary of the program's ups and

downs as recorded in the SAR would complement the statistical data gathered on the program and entered into the spreadsheet. Perhaps the qualitative and quantitative data together would provide early indicators or "useful lessons learned" data.

When we considered developing an advanced relational database, we thought this written program description could contain a key word/phrase component that would allow searching the database by single element or a combination of a few elements/key words. The attached C-17 program written characterization is an example.

However, it was determined the added effort, particularly without the advanced database, was not warranted.

C-17 Program Characterization

Profile

EMD Duration:	10 yrs 9 mos	M/S II date: Feb. 85
EMD Cost Overrun:	38 percent	M/S III date: Nov. 95
EMD Schedule Overrun:	79 percent	

Performance Success Ratings	AFOTEC	DOT&E
Operational Effectiveness:	5	5
Operational Suitability:	5	5
Overall Performance:	5	5

(5 is very good, max. down to 1, poor, min.)

Technical Aspects

The C-17 technical development effort was singularly successful. At the beginning of EMD the required technical characteristics were limited to two technical and five operational; all related to major required characteristics of large transport aircraft. Two years later the USD(A&T) directed the inclusion of four additional requirements, and five years into EMD two additional requirements were added. Nine years into EMD an additional four requirements were added, totaling 17 critical requirements. However, the final requirements defined in the ORD and validated by the Joint Requirements Oversight Council (JROC) were only seven, again all related to the major intended use of this cargo aircraft, including paratrooper drops.

The major problems encountered during technical development included a static load wing failure, and a case of two paratroopers making contact during a dual door airdrop. Both problems were solved satisfactorily. The Live Fire Test requirements of the law and imposed by DOT&E was a protracted and contentious issue. However, in the SAR reporting, the subject is treated in a professional matter-of-fact manner. The addition of a Defensive Avionics Capability to the aircraft and a rigorous Live Fire Test and Evaluation (LFT&E) program was successfully completed.

Initially engineering manpower was in short supply, impacting design changes, and overall cost and schedule. The Program Management Office (PMO) authorized several technical initiatives: Product

Variability Reduction, Product Engineering, and Value Engineering. And in the last year of EMD a Product Enhancement/Performance Improvement task was awarded to "mature the aircraft."

The overall technical design effort resulted in an aircraft that closely met all its original requirements. Although a few minor limitations were noted in the IOT&E reports, both reports were extremely positive. The design team has also been awarded several prestigious industry awards for their efforts.

Business Aspects

The business aspects of the C-17 program were more tumultuous than the technical. The initial EMD annual SAR report talks of Government reviews rating the

Contractor "best yet" of any aircraft manufacturer, and "ready for production" in a later SAR. These initial reports quickly changed as cost, schedule and quality problems began to surface. These, in turn, generated considerable outside interest and direction for the program office. Involved groups included Congressional direction, Secretary of Defense direction, a Joint DoD Cost/Schedule Control Systems Criteria (C/SCSC) Surveillance Review, and establishment of a Senior Management Review Group in OSD, and advice and consultation by the Defense Science Board (DSB). Extensive Contractor reorganization and management disruptions exacerbated the program's problems.

Some directed program requirements resulting from outside interests included changes to the way costs were being recorded, cuts in FY90 and 91 procurement funds, a 90 percent reduction in planned procurement quantity, and restrictions on the spending of procurement funds until certain events occurred on the production line and certain operational assessments were completed. By far the most serious business impact upon the program was the 1993 Defense Acquisition Board (DAB) decision to only commit to buy 40 aircraft (A/C) prior to the M/S IIIB Full Rate Production decision scheduled for November 1995. A 40 A/C buy was considered adequate to evaluate whether cost, schedule and performance of the C-17 warranted the full buy of 120. Concurrently, the DoD initiated a study of alternative A/C procurement buys to fulfill the mission. Integrated Product Teams (IPTs) were also initiated at this time.

Notwithstanding this hectic business activity and early program EMD problems, the C-17 program completed EMD in a superior manner. Beginning in 1994 the SARs were able to report the timeliness of A/C delivery has "markedly improved" and quality rework was down 40 percent. The C-17 aircraft set at least 22 world performance records by this time, and the Government-Industry team was given several awards. The most recent award (1998) was the Malcolm Baldrige National Quality Award to the (now) Boeing Airlift and Tanker Team, considered one of the very best in achieving quality and world-class business performance.

Analysis

The C-17 program had its share of technical, and in particular, management problems. From a rocky start it emerged from EMD as an acknowledged success. Possible reasons for this include a

limited set of stated technical requirements, the impact of outside guidance and threat of considering an alternate aircraft to fulfill the mission, and time to incorporate all these events within EMD, prior to full rate production and operational use. The EMD phase stretched from a planned 6 years to near 11 actual years.

Key Phrases

Technical Requirements	(clarity)
External Influences	(beneficial)
EMD Duration	(sufficient)

15. The Right Stuff

Beginning in 1990, under the direction of Dr. Owen Gadeken at DSMC, a major research effort was conducted by interviewing 74 program managers and by surveying an additional 356 program management personnel. The objective of the research was to identify the characteristics that make excellent program managers. The results are the eight characteristics shown below. The underlined were identified as being of great importance.

- <u>Are strongly committed to a clear mission</u>
- Have a long term and big picture perspective
- Are both systematic and innovative thinkers
- Find and empower the best people for their product teams
- Are selective in their involvement in program issues
- Focus on external stakeholders
- <u>Thrive on relationships and influence</u>
- Proactively gather information and insist on results

This research has recently been compared with a more recent "Armstrong Thesis" of 1999. The following chart shows the comparative values of similar attributes measured nine years later. Note that three of the four highest-ranked competencies in both research studies are the same.

The above is simply an example of the excellent acquisition research that exists. This type data combined with the DSMC EMD research project quantitative data should be very useful in providing a "template" against which a program's M/S II EMD plan can be compared. If the current Excel spreadsheet of the DSMC research was replaced by a relational database, these comparisons and aggregations of positive attributes could be facilitated.

Program Manager's Leadership Competencies – Historical Comparison

1990 DSMC Study (128 PMs)	1999 Armstrong Thesis (39 PMs)
1. Sense of ownership/mission	1. Long-term perspective
2. Long-term Perspective	2. Innovativeness
3. Managerial orientation (16)	3. Political awareness
4. Political awareness	4. Sense of ownership/mission
5. Optimizing (20)	5. Relationship Development
6. Results orientation	5. Action orientation
7. Innovativeness	7. Strategic influence (13)
7. Systematic thinking (14)	8. Results orientation
9. Focus on excellence	9. Focus on excellence
10. Relationship development	10. Collaborative influence (18)
11. Action orientation	11. Professional expertise (21)
12. Coaches others	12. Coaches others

Repeat numbers = Tie in rankings

( ) = Rank in other survey

Bibliography

Gailey, C. K., Reig, R. W., Weber, W., and George, L. R., *A Study of the Relationship Between Initial Production Test Articles Used in a System Development Program and the Success of the Program,* Technical Report 2-95, May 1995, Defense Systems Management College, Fort Belvoir, VA.

Reig, R. W., Gailey, C. K., Swank, W. J. and Alfieri, P. A. *Department of Defense Acquisition Management Metrics,* Technical Report TR 1-99, October 1999, Defense Systems Management College, Fort Belvoir, VA.

Reig, R. W., Baselining Acquisition Reform, *Acquisition Review Quarterly*, Vol. 7, No. 1, Winter 2000, Defense Systems Management College, Fort Belvoir, VA.

Management and Measurement, *Program Manager*, Vol. 28, No. 3, May-June 1999

Successes and Failures in the last 10 years of the DoD Acquisition System, *Program Manager*, Vol. 24, No. 1, January-February 1995.

Insufficiently Robust DT&E Means Trouble Ahead for OT&E, Program Manager, Vol. 23, No. 4, July- August 1994

Test and Evaluation Guide, March 1998, Defense Systems Management College, Fort Belvoir, VA.

Risk Management Guide, May 1999, Defense Systems Management College, Fort Belvoir, VA

Systems Engineering Fundamentals, Supplementary Text, December 1999, Defense Systems Management College, Fort Belvoir, VA.

Authors' Notes

Major William J. Swank, USAF, is a Professor of engineering management at DSMC. He is a graduate of Ohio Northern University, Troy State University, the Air Command and Staff College, and the DSMC Program Management Course. He is a member of the DSMC Test and Evaluation Department and has experience in weapon systems development, test and evaluation, and aircraft maintenance.

Dr. Paul A. Alfieri is a Professor of engineering management at DSMC. He is a graduate of Virginia Tech, The George Washington University, The Naval Postgraduate School, The United States Naval Academy, and the DSMC Program Management Course. Dr. Alfieri was the third chairman of the Test and Evaluation Department and is now the Course Director for the Advanced Test and Evaluation Course. He has more than 30 years' experience in Navy fleet operations, weapon systems development (aircraft and missiles), program management, and acquisition, fielding and support of Navy materiel. Dr. Alfieri is a retired Commander, U.S. Navy.

Charles K. (C.K.) Gailey III was a Professor of engineering management at DSMC and is now a visiting research professor. He is a graduate of Rice University, Florida Institute of Technology, The Army War College and the DSMC Program Management Course. He was a member of the DSMC Test and Evaluation Department and is now in the College Research Division. He has more than 25 years' experience in the acquisition, fielding and support of Army materiel. Mr. Gailey is a retired Lieutenant Colonel, U.S. Army.

Raymond W. Reig was a Professor of engineering management at DSMC until his retirement and is now a visiting research professor at DSMC. He is a graduate of the U.S. Naval Academy, New York University, Auburn University, Southern Illinois University at Edwardsville, the Air War College and the DSMC Program Management Course. He was the first chairman of the Test and Evaluation Department and is now in the College Research Division. His more than 35 years of experience spans military, government, and private aerospace industry. Mr. Reig is a retired Colonel, U.S. Air Force.